## Supplement for: Nanopore sequencing of the phi X 174 genome

Andrew H. Laszlo<sup>1</sup>, Ian M. Derrington<sup>1</sup>, Brian C. Ross<sup>1</sup>, Henry Brinkerhoff<sup>1</sup>, Andrew Adey<sup>2</sup>, Ian C. Nova<sup>1</sup>, Jonathan M. Craig<sup>1</sup>, Kyle W. Langford<sup>1</sup>, Jenny Mae Samson<sup>1</sup>, Riza Daza<sup>2</sup>, Kenji Doering<sup>1</sup>, Jay Shendure<sup>2</sup>, Jens H. Gundlach<sup>1,†</sup>

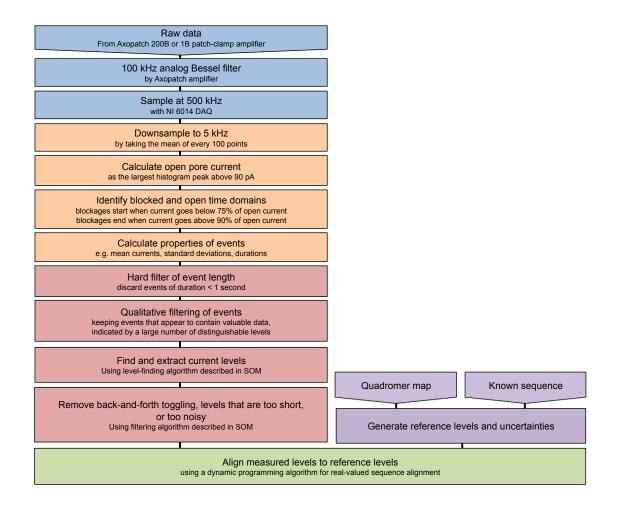
## **Table of Contents:**

Data reduction flow chart (Sup. Fig. 1)	2
Current consensus for all de Bruijn strands (Sup. Figs. 2-4)	3-5
Description of automatic level finding algorithm	6-7
Description of alignment algorithm (Sup. Fig. 5)	8
Alignment of nanopore reads to quadromer prediction (Sup. Fig. 6)	9
Calculating alignment significance (Sup. Fig. 7)	10
Coverage plot for phi X 174 amplicons (Sup. Fig. 8)	11
Full phi X 174 library gel (Sup. Fig. 9)	12
Phi X 174 consensus and quadromer map revision (Sup. Figs. 10, 11)	13-14
DNA scaffold reconstruction (Sup. Fig. 12)	15
Viral alignment and identification (Sup. Fig. 13)	16
SNP calling workflow schematic (Sup. Fig. 14)	17
SNP detection efficiencies and resequencing confusion matrix (Sup. Fig. 15)	18
DNA strands used in this study (Sup. Table 1.1, 1.2)	19-20
Table of quadromer map values (Sup. Table 2.1-2.4)	21-24
References	25

<sup>&</sup>lt;sup>1</sup>Department of Physics, University of Washington, Seattle, WA 98195, USA <sup>2</sup>Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

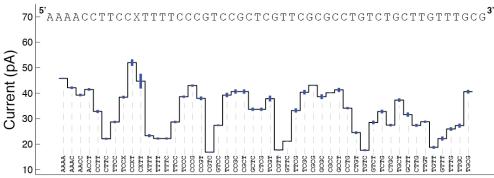
<sup>&</sup>lt;sup>†</sup>Corresponding author. E-mail: <u>gundlach@uw.edu</u>

## Supplementary Figure 1: Data reduction flow chart

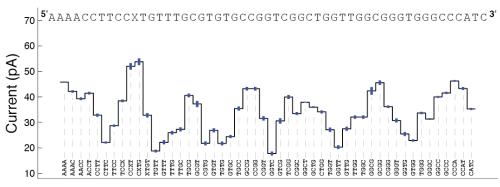


#### Supplementary Figure 2: Current consensuses for all de Bruijn strands 1-3

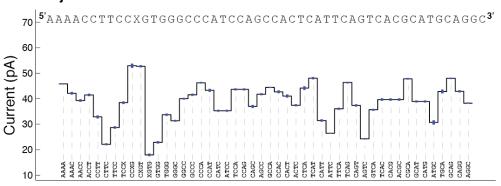
# de Bruijn Section 1



## de Bruijn Section 2



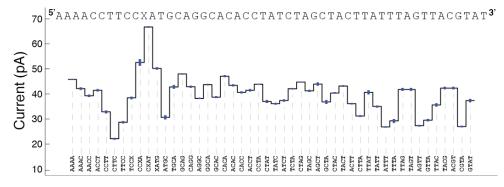
#### de Bruijn Section 3



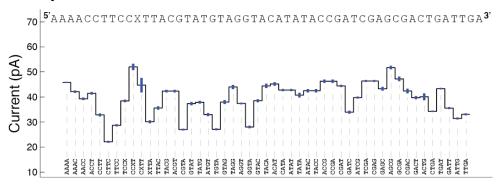
**Supplementary figure 2**: **de Bruijn sequence segments 1-3**: Consensus current level sequences and associated quadromers for de Bruijn segments 1-3. Consensuses for each strand were generated from 17, 11, and 15 reads of strands 1,2, and 3, respectively.

## Supplementary Figure 3: Current consensuses for all de Bruijn strands 4-6

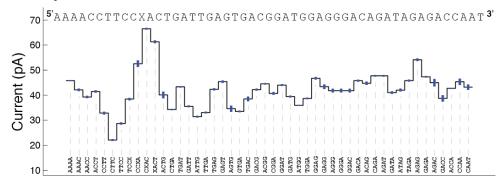
### de Bruijn Section 4



### de Bruijn Section 5



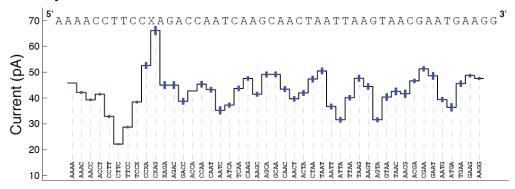
#### de Bruijn Section 6



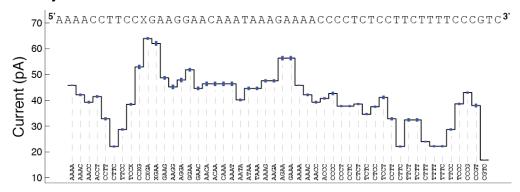
**Supplementary figure 3**: **de Bruijn segments 4-6**:Consensus current level sequences and associated quadromers for de Bruijn segments 4-6. Consensuses for each strand were generated from 12, 11, and 12 reads of strands 4,5, and 6, respectively.

## Supplementary Figure 4: Current consensuses for all de Bruijn strands 7-8

## de Bruijn Section 7



### de Bruijn Section 8



**Supplementary figure 4**: **de Bruijn segments 7-8.** Consensus current level sequences and associated quadromers for de Bruijn segments 7-8. Consensuses for each strand were generated from 14 and 7 reads of strands 7 and 8, respectively.

#### Description of automatic level finding algorithm

Our algorithm identifies levels in two steps. First, it identifies the level boundaries within the time-traces. Then it removes spurious levels and combines levels where the polymerase appears to have "toggled" between two bases.

**Identifying level locations.** Starting from the beginning of a current trace our technique first examines part of a current trace and divides it into two sections. Under the assumption that the sampled currents within each level are Gaussian-distributed, we compute the total probability that the two sections originated from two distinct Gaussian distributions. We divide this total probability by the probability of the null hypothesis, namely that the combination of the two sections originated from a single Gaussian. For ease of computation, we use log probabilities. For a given section the observed mean is  $\hat{I}$  and width is  $\sigma$ . For an individual current measurement within this section, I(t), at time t, the log probability (density) is:

$$\log p(I(t)) = \log \left( \frac{1}{\sqrt{2\pi}\sigma} e^{(I(t)-\widehat{I})^2/2\sigma^2} \right)$$
 Eq. 1.

To find the total log probability that the observations between  $t_1$  and  $t_2$  belong to a single level defined by  $\hat{I}$  and  $\sigma$ , we sum Eq. 1 between  $t_1$  and  $t_2$  and use the definition of sigma, giving

$$\log p(I([t_1, t_2])) = (t_2 - t_1) \log \sigma + \text{const}$$
 Eq. 2.

We calculate total log probability (Eq. 2) for each of the two sections between  $t_1$  and  $t_2$ , and then between  $t_2$  and  $t_3$ . To compare the log probabilities, we subtract the total log probability of the null hypothesis by computing Eq. 2 between  $t_1$  and  $t_3$ . Combining all probabilities yields a comparison metric, which we denote  $\log p(t_1, t_2, t_3)$ ,

$$\log p(t_1, t_2, t_3) = (t_2 - t_1) \log \sigma(t_2, t_1) + (t_3 - t_2) \log \sigma(t_3, t_2) - (t_3 - t_1) \log \sigma(t_3, t_1) + \text{const}$$
Eq. 3.

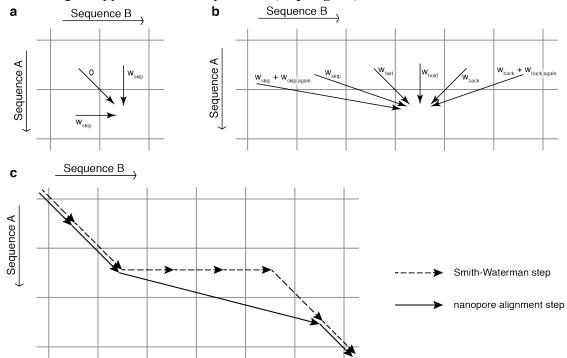
The  $t_2$  yielding minimal  $\log p(t_1, t_2, t_3)$  indicates the location of a possible level transition within the current observations between  $t_1$  and  $t_3$ . In our level finding algorithm, we begin with a given time window ( $[t_1, t_3]$ ) and search for  $t_2$  that minimizes  $\log p$ . If min( $\log p$ ) is less than a specified threshold (we used a threshold of  $\log p = -50$ ) there is a level transition at  $t_2$ , and we recursively search between  $t_1$  and  $t_2$  and between  $t_2$  and  $t_3$  for other level transitions. If min( $\log p$ ) is above the threshold for the original time window then there are no transitions within  $t_1$  and  $t_3$  and we consider a larger window by increasing the value of  $t_3$ .

Not all levels correspond to the single nucleotide forward motion of the phi29 DNAP. We remove levels that (1) have durations  $< 500\mu s$ , (2) have mean values outside of expected ranges (10-70 pA) or (3) have an error in the mean greater than 5 pA. Finally, using the logp calculated using Eq. 3 above, we identify regions of levels that are similar, and combine the levels in them.

#### Description of alignment algorithm

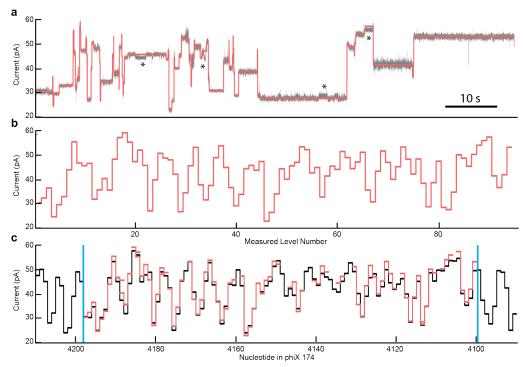
Our tool for aligning level sequences is based on the well-known Needleman-Wunsch and Smith-Waterman algorithms for sequence alignment (1, 2). A Needleman-Wunsch or Smith-Waterman alignment of two base sequences A and B allows for gaps in both sequences. Due to possible gaps the optimal alignment between the first  $n_A$  bases of A and the first  $n_B$  bases of B is one of the following: 1) the optimal alignment between the first  $n_{A-1}$  bases of A and all  $n_B$  bases of B plus a gap in A; 2) the optimal alignment between all  $n_A$  bases of A and the first  $n_{B-1}$  bases of B plus a gap in B; or 3) the optimal alignment between the first  $n_{A-1}$  bases of A and the first  $n_{B-1}$  bases of B plus a final matched base in both sequences. Longer optimal alignments are recursively calculated from shorter optimal sub-alignments, and the entries in the alignment table (sup. fig. 5a) are filled from top-left to bottom-right.

The main complication in our experiment is that our level sequences sometimes step backwards. If we used the Needleman-Wunsch scheme directly each entry in the alignment table would depend on both the entry to the left (due to forward steps) and the entry to the right (due to backsteps), which in turn would depend on the entry in question. To fix this problem, we require each step in the alignment to advance sequence A forward by one, as shown in **Supplementary figure 5b**. Our alignment is therefore the optimal mapping of *every* level in sequence A to its corresponding level in sequence B (or a null level if no good match exists). Note that our alignment trace passes over rather than through skipped levels in sequence B (**Sup. fig. 5c**).



**Supplementary figure 5:** a) Needleman-Wunsch alignments consider horizontal and vertical steps in the alignment table (with a penalty w), corresponding to mismatched bases in one sequence or the other. Diagonal steps, indicating a matched base, generally have no penalty unless there is a mismatch. b) Our nanopore alignment forces every step to progress along sequence A, but allows backsteps in sequence B. We assign affine penalties to skips and backsteps: for example, a backstep of 3 levels would earn the backstep penalty w<sub>back</sub> plus twice the backstep-again penalty w<sub>backagain</sub>. c) The difference between a Needleman-Wunsch alignment and our nanopore alignment when there is a skip in sequence B.

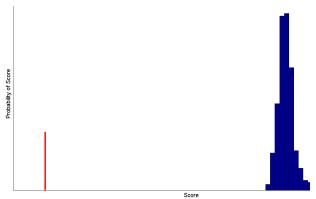
#### Supplementary figure 6: Alignment of nanopore reads to quadromer prediction:



**Supplementary figure 6.** Similar to Fig. 3 in the main text, this figure demonstrates how alignment takes place. Instead of an amplicon being aligned, this figure shows a short sub-segment from a long event. **a**) Our level finding algorithm is used to identify transitions between levels in the current trace. Asterisks mark locations where the algorithm identifies and removes DNAP backsteps. **b**) We then extract the sequence of median current values from each level. **c**) Next, we align the current values to predicted values from the reference sequence using a dynamic programming alignment algorithm similar to Needleman-Wunch alignment (1). In some locations, levels are skipped in the nanopore read either due to motions of the DNAP or errors by the level finding algorithm. We determine read boundaries from the first and last matched levels in the reference sequence. Read boundaries are indicated by the blue lines.

#### Calculating alignment significance

Using the afore mentioned alignment algorithm we align nanopore reads to the predicted level sequence from a known DNA sequence using the quadromer map. The alignment produces a raw score, s, that can be compared to alignments to other reference sequences. Next, we generate a large random sequence along with the expected current levels. We then perform alignments of our measured data to the random sequence. If our measured levels are truly from phi X 174 we expect the score to stand out from the distribution of scores to random alignments. **Supplementary figure 7** shows a histogram of the scores for random alignments (blue) and a marker (red) for the location of the score for the alignment to phi X 174. Strongly negative scores represent good alignments.



**Supplementary figure 7**: The probability distribution of scores for random alignments dP/dS (blue), together with a marker (red) showing the location of the score of the alignment to phi X 174. Units are in arbs. This plot is made using event number 49 from figure 4 of the main

The confidence in the alignment *C* is calculated by

$$C(S_0) = \int_{-\infty}^{S_0} \frac{dP}{dS} dS$$

C represents the probability that a given alignment to a random sequence will produce a score better than  $S_0$ . This particular alignment had a confidence score of  $10^{-147}$ , reflecting a high probability of these measured levels belonging to phi X 174 relative to random alignments. We assume that in the limit of an infinite number of

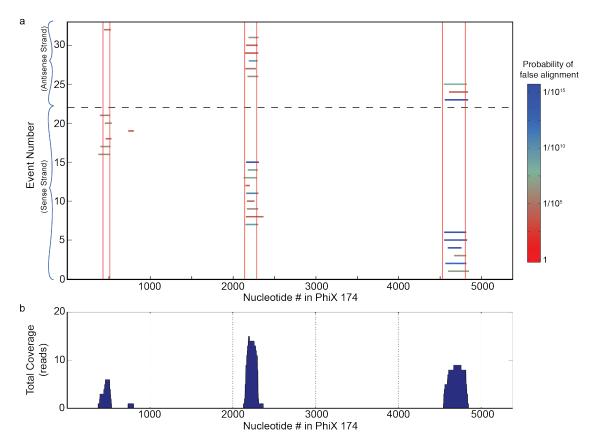
random alignments, the distribution of alignment scores for random sequences approximate a Gaussian, so that

$$C(S_0) = \int\limits_{-\infty}^{S_0} G(S)dS$$

We find the Gaussian by fitting to the width and center of the measured score distribution as floating parameters.

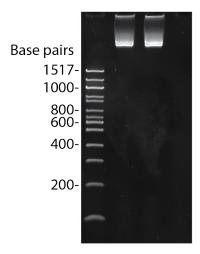
We comment briefly on the meaning of C because of the extreme smallness of these numbers. C represents nothing more than the probability that the produced alignment could also be randomly obtained. The score of  $10^{-147}$  was produced by an alignment of nanopore read of length  $\sim 2000$ .

#### Supplementary figure 8: Coverage plot for phi X 174 amplicons



**Supplementary figure 8: Amplicon alignment.** 31 reads of phi X 174 amplicons aligned to current reference generated by translating the known phi X 174 sequence into current levels using the quadromer map. DNA strands are identified with high confidence, which enables a number of different useful applications such as organism identification and providing a reconstruction scaffold for short high-quality reads obtained with other sequencing technologies. **a)** Alignment bounds for 31 nanopore reads of phi X 174 amplicons. The alignment bounds match well with the actual amplicon locations. All reads with a quality better than 1 in 10<sup>4</sup> fall within one of three locations along the phi X 174 genome revealing the correct location of the amplicons within the genome. Because the adaptors attach to the strands in random orientation, we made reads of both the sense and anti-sense strands. **b)** Coverage for nanopore reads in **a)**.

## Supplementary figure 9: Full phi X 174 library gel



**Supplementary figure 9:** The gel shows the length distribution of phi X 174 genomic DNA used in our experiments (lanes 2 & 3 are replicates). There is a single band that contains broadening towards longer strand lengths. The band broadening is likely due to different numbers of ligated adaptors (0,1, or 2), different adaptor orientations (fan-tail or hairpin), or possibly circular forms of phi X 174. The absence of bands of shorter lengths indicates that our nanopore read lengths are determined not by library quality, but instead by how far the enzyme processes along the DNA before dissociating.

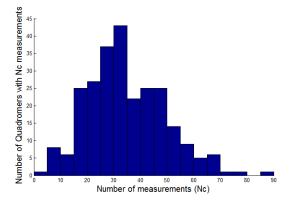
#### Phi X 174 consensus and quadromer map revision

We used values from a current level consensus for the phi X 174 genome to update the quadromer map. To generate the consensus current level sequence for phi X 174, we aligned each nanopore read of the phi X 174 DNA to the predicted current level sequences for its sense and antisense base sequences. The predicted current level sequences were made from the initial measurements of quadromers in the de Bruijn sequence. Alignments with an overall confidence better than  $10^{-6}$  were selected to contribute to the updated map.

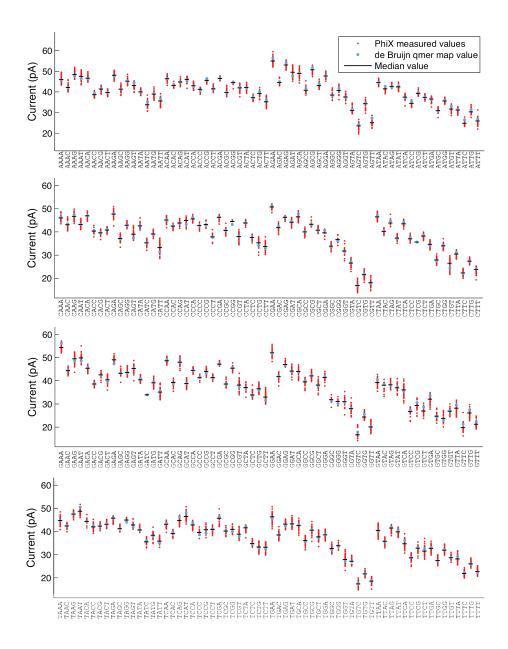
Only levels aligned with high certainty contributed to the consensus. All consensus level values were the average of at least four reads. Also, at least half of all reads covering this level contained a current level matched to that predicted level. The consensus value for the given context was calculated as the mean of all reads aligned to that level.

With the exception of the self complimentary quadromer GATC, there are many instances (35 on average, see **sup. fig. 10**) of each of the remaining 255 quadromers within the 5386-nucleotide phi X 174 genome and its complementary strand. For GATC, we retain the original de Bruijn sequence current value. Using the updated consensus levels, we were able to update the quadromer map with additional measurements in a variety of sequence contexts. The revised map uses the mean and standard deviation of all measurements made of each quadromer throughout the phi X 174 consensus (Table S2). **Supplementary figure 11** (next page) shows the revised quadromer map in comparison to the original quadromer map.

The consensus generation and quadromer map updating procedures were tested by reserving five high-quality reads to be excluded from the generation of the revised quadromer map, and then aligning these reads to both the consensus sequence and the updated prediction. In all cases, the confidence in the alignments of these reads improved dramatically when the new prediction was used as the reference sequence.



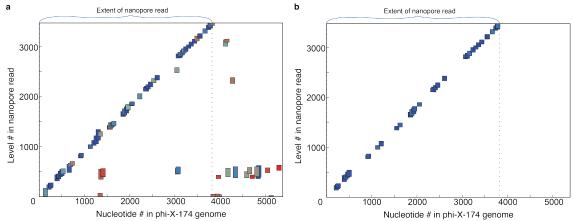
**Supplementary figure 10**: Histogram of the number of instances of each quadromer in the phi X genome. Each quadromer has 35 reads on average.



**Supplementary figure 11:** The revised quadromer map in alphabetical order, with 64 quadromers (written 5'-3') in each panel, beginning with A,C,G,T, respectively. Red dots are measured values of the quadromer in phi X 174. The blue dots are the values from the original de Bruijn quadromer map. The black lines are the medians of all measured instances of a given quadromer.

#### **DNA scaffold reconstruction:**

The difficulty of *de novo* sequencing with most sequencing technologies is that their many short DNA reads must be stitched together in the proper order to form a long contiguous sequence. This assembly process is usually performed by looking for sequence similarity between overlapping reads. We demonstrate an alternative method of sequence scaffolding by mapping 100 short, 100 bp long reads from an Illumina MiSeq sequencer to one of our long (3466 levels) nanopore read. The mapping was performed by converting the sequence of each Illumina read, and its reverse compliment, into a sequence of current levels using our quadromer map, and then using our level alignment tool to find the likely location of the current level sequence in our nanopore read. Figure S13 shows the fate of the 87 (out of 100) Illumina reads which generated an alignment to the nanopore read: 61 Illumina reads lay at least partially within the nanopore read and were aligned properly; 9 Illumina reads lay at least partially within the nanopore read and were misaligned; and 17 Illumina reads fell entirely outside the nanopore read. 9 of the 13 Illumina reads that did not generate an alignment actually lay outside the nanopore read.

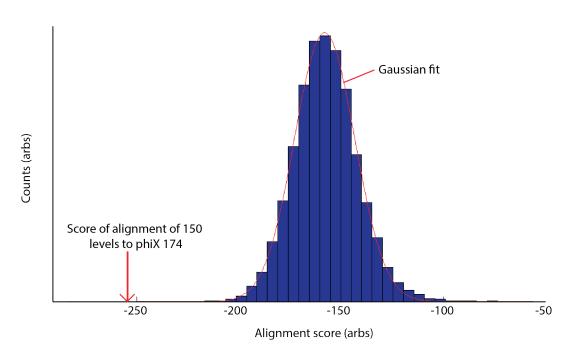


**Supplementary figure 12**: Alignment-scaffolded assembly of 87 short DNA sequences. Each short DNA sequence is indicated by a box, whose horizontal width indicates the location of the Illumina read within the phi X 174 genome and whose vertical height indicates the span of the Illumina read alignment to a 3466 level nanopore read (spanning 3819 bp in the phi X 174 genome). a) Location of all 87 reads that produced alignments to the nanopore read. Color indicates the alignment quality: blue is high-quality and red is low-quality. Overlapping rectangles represent contigs. b) After applying a cutoff filter on nanopore alignment quality and the length of the alignment to the nanopore read (keeping only alignments spanning less than 130 levels) we see that all erroneous alignments are filtered out (plus 23 low scoring but correct alignments). Of the 74 Illumina reads which should have aligned to our nanopore read, we are left with 38 (51%) Illumina reads properly localized within the phi X 174 genome with high confidence.

#### Viral alignment and identification:

Viral identification was performed by aligning arbitrary 250 level subsets of nanopore reads of phi X 174 DNA to a viral database consisting of 5287 viruses (including phi X 174) totaling 156 megabases. To increase the alignment speed, first an 80-level subset of the nanopore read was aligned to every viral genome in the database. This initial alignment was used to generate a list of likely candidate alignments. Alignment confidences for each 80-level alignment to each virus were tallied and compared. Viruses with log confidences better than the mean log confidence score by 3 standard deviations were passed on to the next round; all others were discarded. In the next alignment round, 150 levels of the nanopore read were aligned to the remaining viruses followed by another round of database reduction. Finally all 250 levels were aligned to the remaining viruses. For each event tested, the 250-level alignment correctly identified the DNA as belonging to phi X 174 with at least 99.9996% confidence (in all instances, the 80 or 150-level alignment also suggested phi X 174 as the most likely although with reduced confidence).

Performing a final alignment of the entire >1000-level nanopore reads to the phi X 174 genome can confirm the conclusion to almost arbitrarily high confidence (less than 1 in  $10^{70}$  chance of mis-identification).



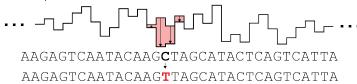
**Supplementary figure 13:** Distribution of alignment scores for a 150-level segment from a long nanopore read to a viral genome database. The distribution of scores is Gaussian. Here the 150-level alignment to phi X 174 differs from random alignments by  $\sim$ 6.5 standard deviations, identifying the strand with 99.99997% confidence.

#### SNP calling workflow schematic

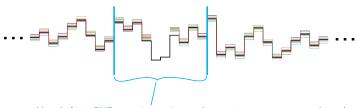
1) We began with measured reference consensus made from several phi X 174 reads



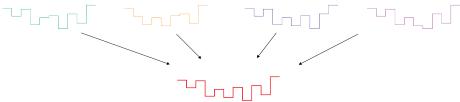
2) We inserted fake SNP's by altering sections of the reference consensus with quadromer map values for the SNP. In this illustration, a  $\mathbf{C}$  is replaced with a  $\mathbf{T}$ .



3) We then aligned several nanopore reads to the modified consensus. In general reads aligned quite well to the consensus, alignment errors may occur near inserted SNP's. We used alignments to identify the region of the nanopore reads that will be scrutinized for making the SNP call.



4) We extracted measured levels from SNP-covering region and generate a consensus using a local consensus generating algorithm which aligns multiple sequences to one another and generates a consensus.

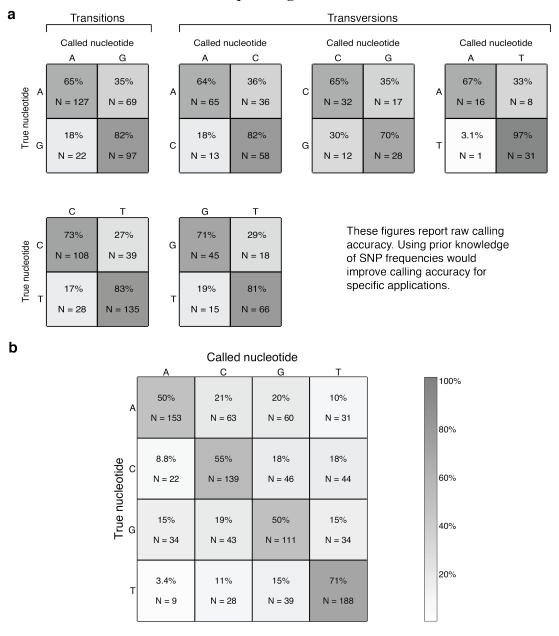


5) Finally we aligned the consensus to the two different SNP possibilities and made a call. Alignments to incorrect sequence resulted in errors (skips, backsteps, holds, bad levels) which decreased the quality of the alignment score. The DNA sequence that matched best with the consensus was called as the measured nucleotide. Including prior probabilities for allele frequency can be used to increase calling accuracy.



**Supplementary figure 14**: Schematic outline of SNP detection methods. We inserted "mock SNPs" into a reference consensus by inserting quadromer map values corresponding to the inserted SNP. Transversions and transitions were inserted into the genome in the following ratio (70% C<->T/G<->A, 15% C<->A/G<->T, 9% G<->C, and 5% A<->T) corresponding to how often they occur within the human genome (3). We then performed alignments of nanopore reads to the reference consensus as if we were comparing new nanopore reads to a previously measured consensus. We used these alignments to extract current levels from events that had reads of the SNP region in question. We then generate a consensus using these nanopore measurements. The sequence that aligns best with the consensus is selected as the measured allele. See **Sup. Fig. 16** for detection efficiencies.

#### SNP detection efficiencies and resequencing confusion matrix



**Supplementary figure 15: Confusion matricies for SNPs and reference sequencing. a)** Detection efficiencies for each possible SNP in each box. The actual DNA nucleotide is displayed along the left of each box while the nanopore call is displayed along the top of each box. The contrast within the box indicates our ability to distinguish between the two nucleotides in various sequence contexts. **b)** shows detection efficiency for reference sequencing where instead of comparing only two nucleotides (as one does when interrogating most SNPs), we select the nucleotide that matches the data best out of all four nucleotides. Combining reads of both sense and anti-sense strands can increase calling accuracy.

## **Supplementary Table 1.1: DNA strands used in this study:**

Construct	Sequence	# of Pores	# of Events
deBruijn Section 1	5' PABABABACCTICCXTITICCCGTCCGCTCGTTCGCCCTGTTCTTTGCGTCGTCCTCCT	4	7
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose	(Calibration) (Blocker) (Brimer)		
deBruijn Section 1 Blocker	5' TTAGCCGACCGGCACGCAAACAAGCAGCAXXXXXXX 3'		
Nucleotide number	0 5 10 15 20 25 30 35 40		
deBruijn Section 2	5' PAAAAAAACCTICCXIGITIGCGIGGCGGIGGGCIGGTIGGCGGGGGCCCCAICAAAACACTCAIAAGCAITCICAIGCAGGICGIAGCC 3'	7	14
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose	(Calibration) () (Bquence		
deBruijn Section 2 Blocker	5' TIATGAGTGTTTTGATGGGCCCACCCGCCAXXXXXXX 3'		
Nucleotide number	0 5 10 15 20 25 30 35 40		
C noite Continue	C   2007 AREA DESCRIPTION OF A REPORT ARTHUR ARTH	u	13
Nicleotide number		ר	71
DNA purpose	(Calibration-) () (Blocker-) (Pri		
deBruijn Section 3 Blocker	GTGACTXXXXX		
Nucleotide number	0 5 10 15 20 25 30 35 40		
deBruiin Section 4	5. PAAAAAAACCTICCXATGCAGGCACACCTATCTAGCTACTTAGTTAGTTAGTTAG	5	11
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose	(Calibration) (		
deBruijn Section 4 Blocker	5' TIGTALATIGTACCTACATACGTAACTAAATAAXXXXXXXX 3'		
Nucleotide number	0 5 10 15 20 25 30 35 40		
Legend			
X = abasic site			
Z = 3-carbon spacer			
P = phosphate group			

## **Supplementary Table 1.2: DNA strands used in this study:**

Construct	Sequence	# of Pores	# of Events
deBruijn Section 5	S' PABABABACCTICCXITACGIBIGIBGGIBCAIBIBACGGAICGAGGGACTGAITGAGTGACGGAIGGAGGAIGCAGTCICAIGCAGGICGIBGCC 3'	9	12
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose	(Calibration) (		
deBruijn Section 5 Blocker	ATCCGTCACTCAATCAGTCGCTCGAT		
Nucleotide number	0 5 10 15 20 25 30 35 40		
deBruijn Section 6	5. PAAAAAAACCTTCCXACTGATTGAGTGACGGATGGAGGGACAGATAGAGACCAATCAAAGAAAAAAAGCATTCTCAATGCAGGTCGTAGCC 3'	ю	15
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose	(Calibration) (		
deBruijn Section 6 Blocker	S' TITIAITAGIIGCIIGAIIXXXXXXZ 3'		
Nucleotide number	0 5 10 15 20 25		
deBruijn Section 7	5. PARARARACCTICCXAGACCATCAAGGAACTAATTAAGTAAGGAAGGAAGAAATAAAGAAAAACCAAAGCAAAGCAAAGCAAGGTGGAGGTGGTAGCC 3'	8	11
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose			
deBruijn Section 7 Blocker	S' TIGGITITCITIAITIGITCCITCAITGGITACITAXXXXXXX 3'		
Nucleotide number	0 5 10 15 20 25 30 35 40 45		
deBruijn Section 8	5' PABABABACCTICCXGAAGGAACAATAAAGAAAACCCCTCTCTTTTTCCCGTCCGCTCGCT	1	17
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90		
DNA purpose	(Calibration) (		
deBruijn Section 8 Blocker	5' ITCGCGAACGAGGGXXXXXXXZ 3'		
Nucleotide number	0 5 10 15 20 25		
deBruijn Primer	GCGTACGCCTACGGTTTTCCGTAGGCGTACGCGGCTACGACCTGCATGAGAA		
Nucleotide number	0 5 10 15 20 25 30 35 40 45 50 55		
Phi X Adaptor	S' PZABABABACCTTCCZCCTTCCCATCATCATCAGATCTCACGGG 3'		
Nucleotide number	0 5 10 15 20 25 30 35 40 45		
Phi X Hairpin (1)	5' PCGCCIACGGITITICCGTAGGCGTACGCUTACTIGITACTIGGGG 3'		
Nucleotide number	0 5 10 15 20 25 30 35 40 45		
Phi X Hairpin (2)	5' CCGCCAAGTAAGAGGTA 3'		
Nucleotide number	0 5 10 15 20 25		
Legend			
X = abasic site			
Z = 3-carbon spacer			
P = phosphate group			

# Supplementary Table 2.1: Table of quadromer map values beginning with A.

	de Bruijn		Error on Revised	Number of Phi X 174
Quadromer	Value (pA)	Revised Value(pA)	Value (pA)	Measurements
AAAA	45.8	46.2	1.4	65
AAAC	42.1	42.1	1.3	45
AAAG	47.5	48.5	1.4	34
AAAT	46.4	47.3	1.7	52
AACA	46.4	46.8	1.1	27
AACC	39.3	38.7	0.8	45
AACG	41.6	41.4	0.9	42
AACT	39.7	39.7	1.0	57
AAGA	47.5	48.0	1.2	42
AAGC	41.5	41.4	1.2	25
AAGG	45.2	45.2	1.2	25
AAGT	44.4	43.0	1.6	36
AATA	40.2	40.1	0.9	44
AATC	35.2	34.1	1.5	34
AATG	39.4	38.7	1.4	29
AATT	36.7	35.4	1.8	54
ACAA	46.4	46.4	1.3	34
ACAC	43.4	43.1	0.8	15
ACAG	44.7	44.8	1.0	23
ACAT	45.2	45.9	1.4	16
ACCA	42.7	43.0	1.0	50
ACCC	40.7	41.0	0.9	14
ACCG	46.3	45.4	0.7	40
ACCT	41.5	41.5	1.0	34
ACGA	46.6	46.5	0.6	48
ACGC	39.7	39.9	1.3	31
ACGG	44.5	44.4	0.8	26
ACGT	42.3	41.6	1.5	30
ACTA	41.9	42.1	1.3	43
ACTC	37.4	37.1	0.8	41
ACTG	40.1	39.3	1.4	42
	36.1			48
ACTT AGAA		35.3	1.7	48
	56.3	55.1	1.6	
AGAC	45.0	44.6	1.2	36
AGAG	54.1	53.0	1.3	29
AGAT	47.7	49.3	2.2	19
AGCA	49.1	48.9	1.5	26
AGCC	41.8	41.0	1.4	26
AGCG	51.7	50.4	1.4	26
AGCT	44.0	43.1	1.1	20
AGGA	47.9	47.7	1.1	26
AGGC	38.3	38.8	1.4	13
AGGG	41.8	40.6	2.0	11
AGGT	37.4	37.6	1.3	39
AGTA	31.5	31.4	1.4	53
AGTC	24.3	23.3	1.8	50
AGTG	34.6	33.8	1.9	30
AGTT	27.3	25.2	1.4	60
ATAA	44.6	44.5	0.9	37
ATAC	42.5	41.5	1.0	33
ATAG	42.0	42.8	0.8	41
ATAT	42.8	42.3	0.9	16
ATCA	37.2	37.5	1.1	20
ATCC	35.3	34.4	1.1	17
ATCG	39.8	39.3	1.0	28
ATCT	37.4	37.3	0.9	26
ATGA	36.4	36.4	1.1	25
ATGC	30.7	31.0	1.0	25
ATGG	36.0	35.6	0.8	29
ATGT	33.0	31.7	1.4	14
ATTA	31.6	31.1	0.9	46
ATTC	26.4	25.0	1.0	33
ATTG	31.4	30.6	1.2	39

# Supplementary Table 2.2: Table of quadromer map values beginning with C.

Quadromer	de Bruijn Value(pA)	Revised Value(pA)	Error on Revised Value (pA)	Number of Phi X 174 Measurements
CAAA	46.4	46.1	1.1	36
CAAC	43.5	43.1	1.4	41
CAAG	47.5	46.7	1.3	30
CAAT	43.2	43.3	1.2	31
CACA	47.1	46.9	1.3	15
CACC	40.6	40.3	1.1	20
CACG	39.7	39.4	1.0	18
CACT	41.1	40.8	0.8	20
CAGA	47.7	47.8	1.5	22
CAGC	37.0	37.1	1.2	31
CAGG	42.9	43.0	0.8	18
CAGT	37.4	39.1	1.7	51
CATA	42.8	42.6	1.0	31
CATC	35.3	35.1	1.3	25
CATG	38.9	39.4	1.0	18
CATT	31.4	33.1	2.0	34
CCAA	45.3	45.1	1.0	34
CCAC	42.7	42.3	0.7	15
CCAG	43.7	44.0	1.1	29
CCAT	43.3	44.5	1.3	36
CCCA	46.3	45.5	1.1	15
CCCC	42.6	42.4	1.2	10
CCCG	43.0	43.3	1.2	7
CCCT	37.8	38.0	1.4	19
CCGA	46.3	46.2	0.9	28
CCGC	40.7	40.7	1.1	40
CCGG	43.2	44.2	0.8	19
CCGT	38.0	37.9	1.9	33
CCTA	43.9	43.8	0.9	26
CCTC	37.8	37.4	0.9	27
CCTG	34.1	35.2	1.7	18
CCTT	32.8	33.6	1.7	38
CGAA	51.3	50.6	0.9	58
CGAC	42.4	41.8	1.2	44
CGAG	46.4	46.1	0.7	32
CGAT	44.4	44.1	1.1	19
CGCA	47.8	46.4	1.4	45
CGCC	40.2	39.9	1.0	31
CGCG	43.1	43.5	0.9	26
CGCT	40.7	40.9	0.8	30
CGGA	40.7	39.6	1.3	28
CGGC	33.5	33.8	1.0	25
CGGG	36.2	36.7	1.8	9
CGGT	31.6	31.8	2.0	43
CGTA	27.0	26.2	1.5	41
CGTC	16.8	16.6	1.6	49
CGTG	21.8	21.4	1.4	19
CGTT	17.7	17.9	1.3	58
CTAA	47.3	46.5	1.3	48
CTAC	40.4	40.1	0.9	34
CTAG	44.8	44.4	1.8	1
CTAT	37.0	37.5	1.1	40
CTCA	44.1	43.6	0.9	30
CTCC	37.5	37.2	1.0	27
CTCG	35.7	35.7	2.3	32
CTCT	38.6	38.2	0.9	34
CTGA	34.3	34.5	1.1	24
	27.5	27.9	1.0	38
CTGC			1.0	23
CTGC CTGG	34.1	33.8	1.0	20
		33.8 26.1	2.1	22
CTGG	34.1			
CTGG CTGT	34.1 24.5	26.1 30.3	2.1	22
CTGG CTGT CTTA	34.1 24.5 31.2	26.1	2.1 1.0	22 36

## Supplementary Table 2.3: Table of quadromer map values beginning with G.

	de Bruijn		Error on Revised	Number of Phi X 174
Quadromer	Value (pA)	Revised Value(pA)	Value (pA)	Measurements
GAAA	56.3	54.3	1.3	50
GAAC	44.6	44.1	1.1	44
GAAG	48.7	49.3	1.9	38
GAAT	48.6	50.0	1.6	31
GACA	45.8	45.1	1.2	24
GACC	38.7	38.6	0.9	30
GACG	42.2	42.4	0.9	38
GACT	39.7	40.4	1.5	49
GAGA	47.4	49.1	1.2	23
GAGC	43.3	43.4	1.2	21
GAGG	43.4	43.8	1.3	16
	45.4	45.2		
GAGT			1.8	49
GATA	41.0	40.5	0.9	16
GATC	33.9	33.9	2.3	6
GATG	39.4	39.2	1.3	18
GATT	35.6	35.1	1.9	24
GCAA	49.1	48.8	0.8	41
GCAC	38.7	39.1	1.5	22
GCAG	48.0	47.8	1.4	35
GCAT	38.9	38.8	1.3	21
GCCA	44.4	44.7	0.7	22
GCCC	41.5	41.2	1.4	8
GCCG	43.2	43.9	1.4	41
GCCT	43.2	43.9	0.8	21
GCGA	47.2	47.3	0.6	23
GCGC	38.6	38.8	1.2	22
GCGG	45.6	45.5	1.0	29
GCGT	37.2	38.1	1.9	35
GCTA	36.8	37.1	1.3	21
GCTC	33.7	34.2	1.4	20
GCTG	36.0	36.6	1.3	26
GCTT	31.6	32.9	1.6	33
GGAA	51.8	52.2	1.6	32
GGAC	41.8	41.5	1.6	26
GGAG	46.7	47.4	1.2	27
GGAT	44.1	44.3	1.5	11
GGCA	43.7	43.6	1.7	18
GGCC	40.0	39.3	1.6	8
GGCG	42.3	42.2	1.5	22
GGCT	37.9	38.3	1.8	17
GGGA	41.8	41.3	1.8	14
GGGC	31.3	31.7	1.5	6
GGGG	30.2	31.1	4.6	7
GGGT	30.7	30.9	2.4	15
GGTA	28.0	27.8	1.8	44
GGTC	17.8	17.1	1.8	40
GGTG	25.5	24.7	1.3	24
GGTT	20.3	19.9	1.6	54
GTAA	40.3	39.3	1.7	47
GTAC	38.5	37.5	1.9	31
GTAG	38.0	38.3	1.4	42
GTAT	37.4	37.2	1.7	44
GTCA	35.7	35.9	2.4	34
GTCC	27.4	26.3	2.1	31
GTCG	30.6	29.2	2.0	58
GTCT	28.5	27.1	1.7	46
GTGA	33.5	31.7	2.0	23
			1.8	24
GTGC	24.4	24.9		
GTGG	22.9	24.0	1.9	25
GTGT	26.9	26.9	1.4	21
GTTA	29.5	27.8	1.9	49
GTTC	21.2	20.1	1.8	47
GTTG	27.5	25.9	1.6	45
GTTT	22.2	21.4	1.2	65
		•		

# Supplementary Table 2.4: Table of quadromer map values beginning with T.

Conditioner		de Bruijn		Error on Revised	Number of Phi X 174
TARC 42.5 42.4 0.9 39 17 TARC 42.6 47.5 1.0 28 TART 50.5 48.7 1.2 48 TART 50.5 48.7 1.2 48 TART 50.5 48.7 1.2 48 TART 50.5 42.5 42.0 1.2 41 TARC 42.5 42.5 42.5 1.0 37 TART 43.2 43.2 43.5 1.0 37 TART 43.2 43.2 43.5 1.1 32 TARC 43.8 45.8 45.3 1.1 32 TARC 44.0 44.8 1.0 20 TARC 44.0 30.7 20 TARC 44.0 38.1 38.5 1.0 28 TARC 44.0 39.5 1.3 56 TARC 44.0 39.5 1.3 38.5 1.0 28 TARC 44.0 39.5 1.3 38.5 1.0 28 TARC 44.0 39.7 39.8 31.1 1.0 22 TARC 44.0 39.7 39.8 39.3 1.2 31 TARC 44.0 39.7 39.8 39.3 1.2 31 TARC 44.0 39.7 39.8 39.3 1.2 31 TARC 44.0 39.7 39.0 1.1 23 31 TARC 44.0 39.7 39.0 1.1 30 TARC 44.0 39.7 39.0 1.1 30 TARC 44.0 39.7 39.0 1.1 23 31 TARC 44.0 39.7 39.0 1.1 23 31 TARC 44.0 39.7 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.3 39.0 39.0	Quadromer		Revised Value(pA)		
TAMC	TAAA		44.7	1.2	56
TART	TAAC	42.5	42.4	0.9	39
TACA	TAAG	47.6	47.5	1.0	28
TACC	TAAT	50.5	48.7	1.2	48
TACC	TACA	44.4	44.3	1.4	24
TACT	TACC	42.5	42.0	1.2	41
TACA  TACC  41.3  41.4  1.0  20  TAGG  44.0  44.8  0.7  25  TAGT  41.8  42.9  1.3  56  TATT  40.7  40.9  0.9  34  TATC  36.1  35.5  1.0  26  TATT  TATT  35.0  36.0  1.1  51  TATT  36.1  39.7  39.0  1.1  23  TCAC  39.7  39.0  1.1  23  TCAC  48.1  48.1  46.5  1.4  36  TCCA  43.7  42.6  1.3  TCCC  39.3  40.8  1.0  39.3  40.8  1.1  100  35  TCCC  39.3  40.8  100  39.3  40.8  1.1  100  35  TCCC  40.4  40.7  10.0  35  TCCC  40.4  40.0  1.2  37  TCCC  37.9  38.9  1.4  71  TCTTA  42.1  41.3  11.2  33  TCTTA  42.1  41.3  11.2  33  TCTTCTA  42.1  41.3  11.2  33  TCTTCTA  42.1  43.1  11.4  46.6  10.3  10.7  TCCC  32.8  33.1  1.1  1.3  23  TCCC  32.8  33.1  1.3  23  TCCT  TCCC  32.8  33.1  1.3  23  TCCT  TCCC  32.8  33.1  1.3  23  TCCT  TCCC  34.6  43.1  1.4  46.6  46.6  32.2  TCCC  32.8  33.1  1.3  23  TCTTC  TCTA  42.1  43.3  1.4  46.6  46.0  2.1  31.1  31.1  TCTC  32.4  33.1  1.4  48.6  49.7  TCCC  37.7  TCCC  38.5  38.9  1.4  TCTC  37.7  TCTC  34.6  38.9  1.4  71  TCTC  34.6  34.8  35.1  1.9  32.8  TCCC  38.7  TCCC  39.7  39.0  1.1  1.2  24  TCAC  39.7  TCCC  37.7  TCCC  37.7  TCCC  37.7  TCCC  38.6  1.1  1.2  23  TCCC  39.7  TCCC  30.0  40.0  41.0  41.0  41.0  41.0  42.1  41.3  42.1  41.3  42.2  43.1  44.8  43.1  44.9  TCCC  38.9  38.9  1.4  TCCC  39.9  39.9  34.4  TCCC  39.9  34.4  TCCC  39.9  34.4  TCCC  39.9  39.9  39.9  39.9  39.9  39.9  39.9  39.9  39.9  39.9  39.9  30.	TACG	42.3	42.5	1.0	37
TAGC	TACT	43.2	43.2	0.8	47
TAGG	TAGA	45.8	45.3	1.1	32
TAGT TATA 40.7 40.9 0.9 34 TATC 36.1 38.8 1.0 26 TATC 36.1 38.8 1.0 26 TATC 37.9 38.3 1.2 31 TATT 35.0 36.0 1.1 61 TCAA 43.7 43.0 1.1 23 TCAC 38.7 TCAC 38.7 38.0 1.1 23 TCAC 38.7 TCAC 38.7 38.0 1.1 23 TCAC 38.7 TCAC 38.7 48.0 1.1 23 TCAC 38.7 TCAC 38.7 48.1 36 TCCC 38.6 46.3 44.8 1.3 36 TCCC 48.1 TCCC 48.1 48.1 1.3 36 TCCC 48.1 TCCC 48.1 1.4 36 TCCC 38.6 39.5 1.1 1.1 18 TCCC 39.6 39.3 40.8 1.6 32 TCCC 40.4 40.7 1.0 35 TCCC 40.4 40.0 1.2 37 TCCC 40.4 40.0 1.2 37 TCCC 40.4 40.0 1.5 30 TCCT 37.9 38.9 1.4 71 TCCC 38.6 38.6 38.9 38.1 1.1 32 TCCT 32.8 33.1 1.1 32 TCCT 32.8 33.1 1.1 32 TCCT 32.8 33.1 1.1 32 TCCT 32.8 38.5 38.6 1.1 32 TCCT 32.8 38.7 38.6 1.1 32 TCCC 38.6 46.0 38.5 38.6 1.1 32 TCCC 38.6 46.0 38.7 38.6 1.1 32 TCCC 38.6 38.7 38.6 1.1 32 TCCC 38.4 38.7 38.6 1.1 1.3 1.6 TCCC 38.4 38.7 38.6 1.1 1.1 1.9 TCCC 39.4 42.9 43.1 1.1 1.2 24 TCCC 39.4 42.9 43.1 1.2 24 TCCC 39.4 42.9 43.1 1.2 24 TCCC 39.4 42.9 43.1 1.1 1.2 24 TCCC 39.4 43.1 1.1 1.2 24 TCCC 39.4 43.3 1.3 1.6 40.6 39.9 1.1 1.6 1.7 1.6 1.7 1.7 1.8 1.8 1.8 1.8 1.8 1.9 1.9 1.9 1.9 1.9 1.9 1.9 1.9 1.9 1.9	TAGC	41.3	41.4	1.0	20
TATA TATC TATC TATC TATC TATC TATC TATC		44.0			29
TATC TATC TATC TATC TATC TATC TATC TATC					
TATG TATT 35.0 36.0 1.1 10AA 43.7 43.0 1.1 10AA 43.7 39.0 1.1 30 10AC 39.7 39.0 1.1 23 10CAC 39.7 39.0 1.1 23 10CAC 46.3 44.8 1.3 35 10CAT 48.1 46.5 1.4 36 10CAC 46.3 43.7 42.6 1.3 28 10CCC 38.6 39.5 1.1 18 18 10CCC 38.6 39.5 1.1 18 18 10CCC 39.3 40.0 1.6 32 10CCT 41.2 40.7 1.0 35 10CAC 46.4 45.2 2.0 42 10CAC 42 10CAC 44.4 45.2 2.0 42 10CAC 42 10CAC 44.4 45.2 2.0 42 10CAC 44.1 45.2 2.0 42 10CAC 42 10CAC 42 10CAC 42 10CAC 43 10.0 11.1 13.1 10.1 10.0 10.5 10.0 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0 10.1 10.0					
TATT TCAA 43.7 43.0 1.1 100 TCAC 39.7 39.0 1.1 223 TCAG 46.3 44.8 1.3 35 TCAT 48.1 46.5 1.4 36 TCCA 48.1 44.8 1.3 36 TCCA 48.1 100 48.1 44.8 1.3 36 TCCA 48.1 100 48.					
TCAA 43.7 43.0 1.1 30  TCAC 39.7 39.0 1.1 23  TCAG 46.3 44.8 1.3 35  TCAT 48.1 46.5 1.4 36  TCAT 48.1 46.5 1.4 36  TCCA 43.7 42.6 1.3 28  TCCC 38.6 39.5 1.1 18  TCCC 38.6 39.5 1.1 18  TCCG 39.3 40.8 1.6 32  TCCG 40.4 41.2 40.7 1.0 35  TCGA 46.4 45.2 2.0 42  TCCG 40.4 46.4 45.2 2.0 42  TCCG 40.0 41.0 1.5 30  TCCT 37.9 38.9 1.4 71  TCTA 42.1 41.3 1.2 33  TCTC 37.9 38.9 1.4 77  TCTA 42.1 41.3 1.2 33  TCTC 32.8 33.1 1.3 23  TCTT 32.4 33.1 1.4 68  TCGA 38.5 38.6 1.1 32  TGAA 45.6 46.0 2.1 31  TGAC 38.5 38.6 1.1 32  TGAA 45.6 46.0 2.1 31  TGAC 38.5 38.6 1.1 32  TGAA 45.6 46.0 2.1 31  TGAC 38.5 38.6 1.1 32  TGAA 45.6 46.0 2.1 31  TGAA 42.1 43.1 1.2 24  TCAT 43.3 1.3 1.2 24  TGAA 45.6 46.0 2.1 31  TGAA 42.1 43.1 1.2 24  TGAA 43.3 31.3 1.4 1.4 68  TGAA 44.8 33.9 1.3 1.3 16  TGAA 42.8 42.4 43.1 1.7 28  TGCA 42.8 42.4 1.7 28  TCCC 35.4 35.9 1.6 27  TCCG 32.1 32.5 1.1 21  TCCG 32.1 32.5 1.1 21  TGCG 33.7 37.9 38.9 1.6 27  TCCG 32.1 32.5 1.1 21  TGCG 32.1 32.5 1.1 21  TGCG 33.7 33.9 1.3 34  TGCT 37.3 37.9 1.3 34  TGCT 37.3 37.9 1.3 34  TGCT 37.3 37.9 1.3 34  TGCT 37.3 38.6 1.1 1.7 43  TGCT 37.3 37.9 1.3 34  TGCT 37.4 22.2 22.3 22.5 22.1 22.2 23.3 24.5 24.1 24.1 24.1 24.1 24.1 24.1 24.1 24.1					
TCAC 39.7 39.0 1.1 23 TCAG 46.3 44.8 1.3 35 TCAT 48.1 46.5 1.4 36 TCCA 43.7 42.6 1.3 28 TCCC 38.6 39.5 1.1 18 TCCG 39.3 40.8 1.6 32 TCCC 38.6 39.5 1.1 18 TCCG 39.3 40.8 1.6 32 TCCG 41.2 40.7 1.0 35 TCGA 46.4 45.2 2.0 42 TCGC 40.4 40.0 1.2 37 TCGG 40.0 41.0 1.5 30 TCGT 37.9 38.9 1.4 71 TCTC 32.8 33.1 1.3 1.2 33 TCTC 32.8 33.1 1.3 23 TCTT 32.4 33.1 1.4 68 TCGA 45.6 46.0 2.1 31 TCGA 45.6 46.0 1.1 32 TCGA 45.8 45.8 11.1 1.2 24 TCGC 35.4 35.9 36.6 1.1 32 TCGC 35.4 35.9 36.6 36.1 32 TCGC 35.4 35.9 36.6 36.1 33 TCGC 35.4 35.9 36.6 36.1 33 TCGC 35.4 35.9 36.6 36.1 33 TCGC 35.4 35.9 36.6 36.0 36.0 36.0 36.7 38.6 37.9 38.6 36.1 37 TCGC 35.4 35.9 36.6 37.9 37.9 37.9 37.9 37.9 37.9 37.9 37.9					
TCAG 46.3 44.8 1.3 35  TCAT 48.1 46.5 1.4 36  TCCA 43.7 42.6 1.3 29  TCCC 38.6 39.5 1.1 18  TCCG 39.3 3.9 40.8 1.6 32  TCCT 41.2 40.7 1.0 35  TCCG 40.4 45.2 2.0 42  TCCG 40.4 40.0 1.2 37  TCCG 40.4 40.0 1.5 30  TCCT 37.9 38.9 1.4 77  TCTA 42.1 41.3 1.2 33  TCTC 34.6 34.8 0.9 34  TCTG 32.8 33.1 1.3 1.2 33  TCTT 32.4 33.1 1.4 68  TCAA 45.6 46.0 2.1 31  TCAA 45.6 46.0 2.1 31  TCAA 43.3 43.3 1.3 1.2 24  TCAA 43.3 43.3 1.3 1.2 22  TCAA 43.3 43.3 1.3 1.7 22  TCAA 43.3 43.3 1.3 1.7 22  TCAA 45.6 46.0 2.1 33.1 1.4 66  TCAA 45.6 46.0 2.1 33.1 1.7 32  TCAA 45.6 46.0 2.1 33.1 1.2 32  TCAA 55.6 46.0 2.1 33.1 1.2 32  TCAA 55.7 38.6 1.1 1.7 28  TCCC 35.6 40.6 40.4 1.7 28  TCCC 35.6 40.6 40.4 1.7 43  TCCC 35.6 40.6 40.4 1.7 43  TCCC 37.1 32.2 5 1.1 22  TCCC 37.1 32.5 1.1 32  TCCC 37.1 32.6 1.1 32  TCCC 37.1 32.7 33.8 1.0 33  TTAA 40.1 40.4 1.6 40  TCCC 37.7 21.7 21.7 1.1 19  TCCC 37.7 22.7 22.7 23.6 1.2 23  TTAA 40.1 40.4 1.6 40  TCCC 37.7 22.7 22.7 23.6 1.2 23  TTAA 40.1 40.4 1.6 40  TCCC 37.7 22.7 22.7 23.6 1.2 23  TTAA 40.1 40.4 1.6 40  TCCC 37.7 22.7 22.7 23.6 1.2 23  TTAA 40.1 40.4 1.6 40  TCCC 37.7 22.7 22.7 23.6 1.2 23  TTAA 40.1 40.4 1.6 40  TCCC 37.7 22.7 22.7 23  TTAA 40.1 40.4 40.4 1.6 40  TCCC 37.7 22.7 22.7 23  TTAA 40.1 40.4 40.4 1.6 40  TCCC 37.1 32.4 31.6 1.6 46  TCCC 37.1 32.4 31.6 1.6 46  TCCC 37.1 32.4					
TCAT					
TOCA 43.7 42.6 1.3 28 TOCC 38.6 39.5 1.1 18 TOCG 39.3 40.8 1.6 32 TOCT 41.2 40.7 1.0 35 TOCG 44.2 40.7 1.0 35 TOCG 44.2 40.7 1.0 35 TOCG 40.4 45.2 2.0 42 TOCG 40.4 40.0 1.2 37 TOCG 40.0 41.0 1.5 30 TOCT 37.9 38.9 1.4 71 TOTA 42.1 41.3 1.2 33 TOTC 34.6 34.8 0.9 34 TOCG 32.8 33.1 1.2 33 TOTG 32.8 33.1 1.4 68 TOCG 32.4 33.1 1.4 68 TOCA 38.5 38.6 1.1 32 TOCT 37.6 42.4 43.1 1.2 23 TOCT 37.6 42.1 31 TOCC 38.5 38.6 1.1 32 TOCC 38.5 38.6 1.1 32 TOCC 38.5 38.6 1.1 32 TOCA 42.8 42.4 43.1 1.2 22 TOCA 42.8 42.4 1.7 28 TOCA 58.5 38.6 1.1 1.2 22 TOCA 58.6 38.7 38.6 1.1 1.7 28 TOCC 38.7 38.6 38.9 1.6 27 TOCC 38.7 38.6 1.1 1.7 43 TOCG 38.7 38.6 1.3 27 TOCG 40.6 40.4 1.7 43 TOCG 38.7 38.6 1.3 27 TOCG 32.1 32.5 1.1 21 TOCG 33.7 33.8 1.0 1.9 1.9 1.9 1.9 1.9 1.9 1.9 1.9 1.9 1.9					
TOCC         38.6         39.5         1.1         18           TCCG         39.3         40.8         1.6         32           TCCT         41.2         40.7         1.0         35           TCGA         46.4         45.2         2.0         42           TCGG         40.4         40.0         1.2         37           TCGG         40.0         41.0         1.5         30           TCGT         37.9         38.9         1.4         71           TCTA         42.1         41.3         1.2         33           TCTC         34.6         34.8         0.9         34           TCTG         32.9         33.1         1.3         23           TCTT         32.4         33.1         1.4         68           TCAA         45.6         46.0         2.1         31           TGAA         45.6         46.0         2.1         31           TGAA         45.6         46.0         2.1         31           TGAG         42.4         43.1         1.2         24           TGAT         43.3         43.3         1.3         1.6           TCGA </td <td></td> <td></td> <td></td> <td></td> <td></td>					
TCCG 39.3 40.8 1.6 32  TCCT 41.2 40.7 1.0 35  TCGA 46.4 45.2 2.0 42  TCGC 40.4 40.0 1.2 37  TCGG 40.0 41.0 1.5 30  TCGT 37.9 38.9 1.4 71  TCTA 42.1 41.3 1.2 33  TCTC 34.6 34.8 0.9 34  TCTG 32.9 33.1 1.3 23  TCTT 32.4 33.1 1.4 68  TCAA 45.6 46.0 2.1 31  TCAC 38.5 38.6 1.1 32  TGAG 42.4 43.1 1.2 22  TGAG 42.8 42.4 1.7 28  TCCA 42.8 42.4 1.7 28  TCCC 38.4 35.9 1.6 27  TCCC 37.3 37.9 38.6 1.3 34  TCTC 37.3 37.9 38.6 1.1 1.2 22  TGAG 42.8 42.4 1.7 28  TCCC 38.5 1.1 22 28  TCCC 38.5 1.1 22 28  TCCC 38.6 1.1 1.2 28  TCCC 42.8 42.4 1.7 28  TCCC 38.1 35.9 1.6 27  TCCC 37.3 37.3 37.9 1.3 34  TCCT 37.3 37.9 1.3 34  TCCT 37.3 38.6 1.1 21  TCCC 37.1 28.0 1.7 43  TCCC 37.1 28.0 1.7 64  TCTA 27.1 27.0 1.6 27  TCTA 27.1 27.0 1.6 27  TCTA 35.4 35.9 1.1 1.1 19  TCTA 35.6 35.6 35.6 1.2 53  TTAC 35.6 35.6 1.2 33  TTAC 35.8 28.8 32.8 1.2 30  TTCC 27.3 27.6 1.0 466  TTCT 28.8 28.8 28.8 0.8 55  TTTA 29.2 28.8 28.8 0.8 55					
TCCT 41.2 40.7 1.0 35 TCGA 46.4 45.2 2.0 42 TCGC 40.4 40.0 1.2 37 TCGG 40.0 41.0 1.5 30 TCGT 37.9 38.9 1.4 71 TCTA 42.1 41.3 1.2 33 TCTC 32.8 33.1 1.3 23 TCT 32.4 33.1 1.4 68 TCAA 45.6 46.0 2.1 31 TCTA 45.6 46.0 2.1 31 TCTA 42.4 43.1 1.2 24 TCAA 45.6 46.0 2.1 31 TCGA 42.4 43.1 1.2 24 TCAA 45.6 46.0 2.1 31 TCGA 42.8 42.4 43.1 1.2 24 TCAT TCAT 43.3 43.3 1.3 1.6 TCCC 35.4 35.9 1.6 27 TCCC 37.3 37.7 38.6 1.3 27 TCGA 38.7 38.6 1.3 27 TCGA 38.7 38.6 1.3 27 TCGA 38.7 38.6 1.3 27 TCGC 37.1 27.0 1.6 27 TCGC 17.6 17.1 28.0 1.7 64 TCGA 27.1 27.0 1.6 27 TCGC 17.6 17.1 1.4 32 TCGC 21.7 21.7 1.1 19 TCGT 27.1 28.0 1.7 64 TCGA 27.1 27.0 1.6 27 TCGC 27.1 28.0 1.7 64 TCGA 35.4 35.9 1.6 35 TTAA 40.1 40.4 1.6 39 TTAA 40.6 39.8 1.1 3 53 TTAA 40.1 40.4 1.6 39 TTAA 40.1 40.4 1.6 39 TTAA 40.6 39.8 1.1 53 TTAA 50.0 34.8 1.0 46 TTCC 22.2 21.8 1.1 1.1 58 TTTAA 50.0 34.8 1.2 30 TTCC 25.5 2.2 21.8 1.1 1.1 58					
TCGA 46.4 45.2 2.0 42 TCGC 40.4 40.0 1.2 37 TCGG 40.0 41.0 1.5 30 TCGT 37.9 38.9 1.4 71 TCTA 42.1 41.3 1.2 33 TCTC 34.6 34.8 0.9 34 TCTG 32.8 33.1 1.3 23 TCTT 32.4 33.1 1.4 66 TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 1.2 24 TGAT 43.3 43.3 1.3 1.3 22 TCGC 35.4 35.5 38.6 1.1 32 TCGC 35.5 38.6 1.1 1.2 24 TGAT 43.3 43.3 1.3 1.3 16 TCCC 35.5 1.1 1.2 24 TGAT 43.3 43.3 1.3 1.3 16 TCCC 35.5 1.1 1.7 28 TCCC 35.1 32.6 1.1 1.7 28 TCCC 35.1 32.5 1.6 27 TCCC 35.1 32.5 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1					
TCGC 40.4 40.0 1.2 37 TCGG 40.0 41.0 1.5 30 TCGT 37.9 38.9 1.4 71 TCTA 42.1 41.3 1.2 33 TCTC 34.6 34.8 0.9 34 TCTG 32.8 33.1 1.3 23 TCTT 32.4 33.1 1.4 68 TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 1.2 24 TGAT 43.3 1.3 1.2 25 TGAG 42.4 43.1 1.2 224 TGAT 43.3 3.1 1.2 224 TGAT 43.3 3 3.1 1.2 224 TGAT 43.3 3 3.1 3.3 1.6 TCGC 34.2 4 33.1 1.7 2.2 24 TGAT 43.3 3 3.1 3.3 1.6 TCGC 40.6 40.4 1.7 28 TCGC 35.4 35.9 1.6 27 TCGC 37.3 37.9 1.3 34 TCGT 37.3 37.9 1.3 34 TCGG 33.7 33.8 1.0 1.7 43 TCGT 37.3 37.9 1.3 34 TCGG 33.7 33.8 1.0 1.1 21 TCGG 33.1 32.5 1.1 21 TCGG 33.7 33.8 1.0 1.1 21 TCGT 27.1 28.0 1.7 64 TCGT 27.1 27.0 1.6 27 TCGT 27.1 28.0 1.7 64 TCGT 27.7 1.1 1.1 1.9 TCGT 21.7 21.7 1.1 1.9 TCGT 21.7 21.7 1.1 1.9 TCGT 35.6 35.6 35.6 1.2 53 TTAA 40.1 40.4 1.6 39 TTAA 40.1 40.6 39.8 1.1 51 TCCC 22.7 28.6 1.3 39 TTGC 33.2 32.5 1.6 40 TTCC 28.7 28.6 1.3 39 TTGC 33.2 32.5 1.6 40 TTCC 27.3 24.4 31.6 1.6 46 TTCA 35.6 35.6 1.2 39 TTCC 27.3 22.6 1.3 32.8 1.2 30 TTCC 27.3 22.4 31.6 1.6 46 TTCA 33.1 32.8 1.2 30 TTCC 27.3 22.6 1.0 46 TTCC 27.3 22.6 1.0 46 TTCC 27.3 27.6 1.0 46 TTCC 27.3 22.6 1.0 46 TTCC 28.8 22.2 22.3 3.1 3.3 67 TTTC 22.2 22.2 21.8 1.1 1.5 58					
TCGC 40.0 41.0 1.5 30 TCGT 37.9 38.9 1.4 71 TCTA 42.1 41.3 1.2 33 TCTC 34.6 34.8 0.9 34 TCTG 32.8 33.1 1.3 23 TCTT 32.4 33.1 1.4 68 TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 1.3 16 TCGA 42.8 42.4 1.7 28 TGGC 35.4 35.9 1.6 27 TGGC 40.6 40.4 1.7 28 TGGC 40.6 40.4 1.7 43 TGGT 37.3 37.9 1.3 34 TGGG 33.7 38.6 1.3 27 TGGG 33.7 38.6 1.3 27 TGGG 33.7 38.6 1.3 27 TGGG 32.1 32.5 1.1 21 TGGG 33.7 33.8 1.0 1.3 27 TGGG 32.1 32.5 1.1 21 TGGG 32.1 32.5 1.1 21 TGGG 33.7 33.8 1.0 14 TGGT 27.1 28.0 1.7 64 TGTA 27.1 27.0 1.6 27 TGTC 17.6 17.1 1.4 32 TGTC 17.6 17.1 1.4 32 TGTT 18.8 18.6 1.0 35 TTAA 40.1 40.4 1.6 39 TTAA 40.1 35.6 35.6 1.2 53 TTAG 33.2 32.5 1.6 62 TTAG 33.1 32.8 1.0 40 TTCC 28.7 28.6 1.3 39 TTAG 41.8 41.2 1.3 53 TTAG 41.8 41.2 1.3 53 TTAG 41.8 41.8 41.2 1.3 53 TTAG 41.8 41.8 41.2 1.3 39 TTCC 28.7 28.6 1.3 39 TTCC 27.3 32.4 31.6 1.6 46 TTGA 33.1 32.8 1.2 30 TTCC 27.3 27.6 1.0 46 TTGC 27.3 27.6 1.0 46 TTGT 22.2 21.8 1.1 58 TTTG 22.2 21.8 1.1 58					
TCGT 37.9 38.9 1.4 71 TCTA 42.1 41.3 1.2 33 TCTC 34.6 34.8 0.9 34 TCTG 32.8 33.1 1.3 23 TCTT 32.4 33.1 1.4 68 TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 2 24 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 1.6 1.6 TGCA 42.8 42.4 1.7 28 TGCC 35.4 35.9 1.6 27 TGCC 35.4 35.9 1.6 27 TGCC 37.3 37.9 1.3 34 TGCT 37.3 37.9 1.3 34 TGCT 37.3 37.9 1.3 34 TGGA 38.7 38.6 1.3 27 TGGC 33.1 32.5 1.1 21 TGGC 33.1 32.5 1.1 21 TGGC 33.7 38.6 1.3 27 TGCC 35.4 35.9 1.6 27 TGCT 37.3 37.9 1.3 34 TGCT 37.3 37.9 1.3 34 TGCT 37.3 37.9 1.3 34 TGGT 37.1 28.0 1.7 64 TGGT 27.1 28.0 1.7 64 TGTA 27.1 27.0 1.6 27 TGTC 17.6 17.1 1.1 1.4 32 TGTC 17.6 17.1 1.1 1.4 32 TGTG 21.7 21.7 1.1 1.9 19 TGTT 18.8 18.6 1.0 35 TTAA 40.1 40.4 1.6 39 TTAA 40.1 40.4 1.6 39 TTAC 35.6 35.6 1.2 53 TTAG 41.8 41.2 1.3 53 TTAG 36.0 34.8 1.1 51 TTCC 28.7 28.6 1.3 39 TTCC 33.4 31.9 32.8 1.1 51 TTCC 33.4 31.9 32.8 1.1 51 TTCC 34.7 28.6 1.3 39 TTGC 35.6 1.2 1.3 39 TTGC 33.2 32.5 1.6 62 TTTAT 40.6 39.8 1.1 51 TTCC 28.7 28.6 1.3 39 TTGC 33.1 32.8 1.2 30 TTGC 37.3 27.6 1.0 46 TTGC 27.3 27.6 1.0 46 TTGT 28.8 28.8 28.8 0.8 55 TTTA 29.2 28.3 1.3 67 TTTC 22.2 21.8 11.1 58 TTTG 22.2 21.8 11.1 58					
TCTA					
TCTC 34.6 34.8 0.9 34 TCTG 32.8 33.1 1.3 23 TCTT 32.4 33.1 1.4 68 TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 16 TGCA 42.8 42.4 1.7 28 TGCC 35.4 35.9 1.6 27 TGCC 35.4 35.9 1.6 27 TGCC 35.4 35.9 1.6 27 TGCC 35.4 36.6 1.3 27 TGCC 37.3 37.3 37.9 1.3 34 TGCT 37.3 37.9 1.3 34 TGCC 32.1 32.5 1.1 21 TGCG 33.7 38.6 1.3 27 TGCC 32.1 32.5 1.1 21 TGCC 32.1 32.5 1.1 21 TGCC 32.1 32.5 1.1 21 TGCC 35.4 36.9 1.0 14 TGCT 37.3 37.9 33.8 1.0 14 TGCT 37.1 27.0 1.6 27 TGTC 17.6 17.1 1.4 32 TGTC 17.6 17.1 1.1 1.9 TGTT 18.8 18.6 1.0 35 TTAA 40.1 40.4 1.6 39 TTAC 35.6 35.6 35.6 1.2 53 TTAG 41.8 41.2 1.3 53 TTAG 41.8 41.2 1.3 53 TTAG 33.1 32.5 1.6 40 TTCC 28.7 28.6 1.3 39 TTCC 28.7 28.6 1.3 39 TTCC 32.4 31.6 1.6 40 TTCC 33.4 31.1 32.8 1.1 1.1 51 TTCC 33.4 31.1 32.8 1.1 1.1 51 TTCC 33.4 31.1 32.8 1.1 1.1 51 TTCC 33.4 33.1 32.8 1.0 46 TTCC 37.3 27.6 1.0 46 TTCC 17.6 17.7 28.0 1.0 46 TTCC 37.3 27.6 1.0 46 TTCC 38.4 31.1 32.8 1.2 30 TTCC 38.7 28.6 1.3 39 TTCC 39.4 31.6 1.6 46 TTCC 37.3 27.6 1.0 46 TTCC 37.3 27.6 1.0 46 TTCC 27.3 27.6 1.0 46 TTCC 27.3 27.6 1.0 46 TTCC 28.8 28.8 0.8 55 TTTA 29.2 28.8 1.3 67 TTTC 22.2 21.8 1.1 58 TTTC 22.2 21.8 1.1 58 TTTC 22.2 21.8 1.1 58					
TCTG 32.8 33.1 1.3 23  TCTT 32.4 33.1 1.4 68  TGAA 45.6 46.0 2.1 31  TGAC 38.5 38.6 1.1 32  TGAG 42.4 43.1 1.2 24  TGAT 43.3 43.3 1.3 1.6  TGCA 42.8 42.4 1.7 28  TGCC 35.4 35.9 1.6 27  TGCC 35.4 35.9 1.6 27  TGCC 35.4 38.7 38.6 1.3 27  TGCC 37.3 37.9 1.3 34  TGCT 37.3 37.9 1.3 34  TGCT 37.3 37.9 1.3 27  TGCG 32.1 32.5 1.1 21  TGGG 33.7 33.8 1.0 14  TGGT 27.1 28.0 1.7 64  TGTA 27.1 27.0 1.6 27  TGTC 17.6 17.1 1.4 32  TGTT 18.8 18.6 1.0 35  TTAA 40.1 40.4 1.6 39  TTAC 35.6 35.6 35.6 1.2 53  TTAC 35.6 35.6 35.6 1.2 53  TTAC 35.6 35.6 35.6 1.2 53  TTAC 36.0 34.8 1.0 1.1 51  TTCC 28.7 28.6 1.3 59  TTCC 33.1 32.5 1.1 51  TTCC 35.6 35.6 1.2 53  TTAC 35.6 35.6 1.2 53  TTAC 35.6 35.6 1.2 53  TTAC 35.6 35.6 1.2 39  TTCC 28.7 28.6 1.3 39  TTCC 28.7 28.6 1.3 39  TTCC 33.2 32.5 1.6 40  TTCC 28.7 28.6 1.3 39  TTCC 32.4 31.6 1.6 40  TTCC 32.4 31.6 1.6 46  TTCC 32.4 33.1 32.8 1.2 30  TTCC 32.4 33.1 32.8 1.2 30  TTCC 32.4 33.6 1.6 46  TTCC 32.4 33.6 1.6 1.6 46  TTCC 32.7 32.4 33.6 1.6 1.6 46  TTCC 32.8 32.1 32.8 1.2 30  TTCC 32.4 33.6 1.6 1.6 46  TTCC 32.8 32.1 32.0 1.0 45  TTTCC 32.8 32.1 32.0 1.0 45  TTTCC 32.8 32.1 32.0 1.0 46  TTCC 32.8 32.1 32.0 1.0 46  TTCC 32.8 32.1 32.0 1.0 46  TTCC 32.9 32.1 32.0 1.0 45  TTTCC 32.9 32.1 32.0 1.0 45  TTTCC 32.9 32.1 32.0 1.0 46  TTCC 32.9 32.1 32.0 1.0 45  TTTCC 32.9 32.1 32.0 0.9 63					
TCTT 32.4 33.1 1.4 68 TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 1.6 TGCC 35.4 35.9 1.6 27 TGCC 35.4 35.9 1.6 27 TGCC 35.4 35.9 1.6 27 TGCC 37.3 37.9 1.3 34 TGCT 37.3 37.9 1.3 34 TGCC 32.1 32.5 1.1 21 TGCC 32.1 32.5 1.1 21 TGCC 33.1 32.5 1.1 21 TGCC 37.1 28.0 1.7 64 TGCT 37.3 37.9 1.3 32.5 1.1 21 TGCC 35.1 35.6 1.3 27 TGCC 35.1 55.6 1.1 51 TGCC 35.1 55.6 1.1 51 TGCC 35.1 55.6 1.1 51 TGCC 17.6 17.1 1.4 32 TGTC 17.6 17.1 1.1 1.9 TGTT 18.8 18.6 1.0 35 TTAA 40.1 40.4 1.6 39 TTAC 35.6 35.6 35.6 1.2 53 TTAC 35.6 35.6 35.6 1.2 53 TTAT 40.6 39.8 1.1 51 TTCC 28.7 28.6 1.3 39 TTCC 28.7 28.6 1.3 39 TTCC 32.4 31.6 1.6 40 TTCC 32.4 31.6 1.6 46 TTCC 32.4 31.6 1.6 46 TTCC 27.3 27.6 1.0 45 TTCC 32.4 31.6 1.6 46 TTCC 27.3 27.6 1.0 45 TTCC 27.3 27.6 1.0 45 TTCC 28.8 28.8 28.8 0.8 55 TTTA 29.2 22.1 32.0 1.0 46 TTCC 28.8 28.8 28.8 0.8 55 TTTA 29.2 22.2 21.8 1.1 58 TTTC 22.2 21.8 1.1 58					
TGAA 45.6 46.0 2.1 31 TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 1.6 TGCA 42.8 42.4 1.7 28 TGCC 35.4 35.9 1.6 27 TGCC 35.4 35.9 1.6 27 TGCT 37.3 37.9 1.3 34 TGGA 38.7 38.6 1.3 27 TGGC 32.1 32.5 1.1 21 TGGT 27.1 28.0 1.7 64 TGTA 27.1 27.0 1.6 27 TGTC 17.6 17.1 1.4 32 TGTT 18.8 18.6 1.0 35 TTAA 40.1 40.4 1.6 39 TTAA 40.6 39.8 1.1 51 TTAC 35.6 35.6 35.6 1.2 53 TTAT 40.6 39.8 1.1 51 TTCC 28.7 28.6 1.3 39 TTCC 28.7 28.6 1.3 39 TTCC 33.2 32.6 1.6 40 TTCC 28.7 28.6 1.3 39 TTCC 33.2 32.6 1.0 35 TTAC 35.6 35.6 1.2 53 TTAC 41.8 41.2 1.3 53 TTAC 36.0 34.8 1.6 40 TTCC 28.7 28.6 1.3 39 TTCC 33.2 32.5 1.6 40 TTCC 32.4 31.6 1.6 40 TTCC 28.7 28.6 1.3 39 TTCC 33.2 32.5 1.6 46 TTCG 32.4 31.6 1.6 46 TTCG 32.4 31.6 1.6 46 TTCG 32.1 32.0 1.0 46 TTGG 32.1 32.0 1.0 46 TTGC 22.2 21.8 1.1 58 TTTA 29.2 28.8 28.8 0.8 55 TTTA 29.2 28.8 1.1 58 TTTC 35.9 26.2 0.9 63					
TGAC 38.5 38.6 1.1 32 TGAG 42.4 43.1 1.2 24 TGAT 43.3 43.3 1.3 16 TGCA 42.8 42.4 1.7 28 TGCC 35.4 35.9 1.6 27 TGCG 40.6 40.4 1.7 43 TGGT 37.3 37.9 1.3 34 TGGA 38.7 38.6 1.3 27 TGGG 33.1 32.5 1.1 21 TGGG 33.7 33.8 1.0 14 TGGT 27.1 28.0 1.7 64 TGTA 27.1 27.0 1.6 27 TGTC 17.6 17.1 1.4 32 TGTC 17.6 17.1 1.5 1.5 35 TTAA 40.1 40.4 1.6 39 TTAA 40.1 40.4 1.6 39 TTAA 40.1 40.4 1.6 39 TTAA 41.8 41.2 1.3 53 TTAT 40.6 39.8 1.1 51 TTCC 28.7 28.6 1.3 39 TTCC 33.2 32.5 1.6 35.6 1.2 53 TTAT 40.6 39.8 1.1 51 TTCC 28.7 28.6 1.3 53 TTCC 33.4 31.6 1.6 40 TTCC 28.7 28.6 1.3 39 TTCC 33.2 32.5 1.6 46 TTCG 32.1 32.9 1.0 46 TTCG 32.1 32.9 1.0 46 TTCG 32.1 32.9 1.0 46 TTCG 32.1 32.0 1.0 46 TTCG 32.1 32.0 1.0 46 TTCG 28.8 28.8 28.8 0.8 55 TTTA 29.2 28.8 1.1 58 TTTA 29.2 28.3 1.3 58 TTTTC 22.2 21.8 1.1 51					
TGAG         42.4         43.1         1.2         24           TGAT         43.3         43.3         1.3         16           TGCA         42.8         42.4         1.7         28           TGCC         35.4         35.9         1.6         27           TGCG         40.6         40.4         1.7         43           TGCT         37.3         37.9         1.3         34           TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGC         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAG <td></td> <td></td> <td></td> <td></td> <td></td>					
TGAT         43.3         43.3         1.3         16           TGCA         42.8         42.4         1.7         28           TGCC         35.4         35.9         1.6         27           TGCG         40.6         40.4         1.7         43           TGCT         37.3         37.9         1.3         34           TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAG         35.6         35.6         1.2         53           TTAT         40.6         39.8         1.1         51           TTC <td></td> <td></td> <td></td> <td></td> <td></td>					
TGCA         42.8         42.4         1.7         28           TGCC         35.4         35.9         1.6         27           TGCG         40.6         40.4         1.7         43           TGCT         37.3         37.9         1.3         34           TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         21.7         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         1.9           TGT         18.8         18.6         1.0         35           TTAA <td></td> <td></td> <td></td> <td></td> <td></td>					
TGCC         35.4         35.9         1.6         27           TGCG         40.6         40.4         1.7         43           TGCT         37.3         37.9         1.3         34           TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA <td></td> <td></td> <td></td> <td></td> <td></td>					
TGCG         40.6         40.4         1.7         43           TGCT         37.3         37.9         1.3         34           TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCC         28.7         28.6         1.3         39 <t< td=""><td></td><td></td><td></td><td></td><td></td></t<>					
TGCT         37.3         37.9         1.3         34           TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG <td></td> <td></td> <td></td> <td></td> <td></td>					
TGGA         38.7         38.6         1.3         27           TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA <td></td> <td></td> <td></td> <td></td> <td></td>					
TGGC         32.1         32.5         1.1         21           TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGG <td></td> <td></td> <td></td> <td></td> <td></td>					
TGGG         33.7         33.8         1.0         14           TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAG         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGG         32.1         32.0         1.0         45           TTGG <td></td> <td></td> <td></td> <td></td> <td></td>					
TGGT         27.1         28.0         1.7         64           TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TCC         28.7         28.6         1.3         39           TCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG					
TGTA         27.1         27.0         1.6         27           TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGA         32.1         32.8         1.2         30           TTGG         33.1         32.8         1.2         30           TTGG         32.1         32.0         1.0         46           TTGT <td></td> <td></td> <td></td> <td></td> <td></td>					
TGTC         17.6         17.1         1.4         32           TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC <td></td> <td></td> <td></td> <td></td> <td></td>					
TGTG         21.7         21.7         1.1         19           TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG <td></td> <td></td> <td></td> <td></td> <td></td>					
TGTT         18.8         18.6         1.0         35           TTAA         40.1         40.4         1.6         39           TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63					
TTAA 40.1 40.4 1.6 39  TTAC 35.6 35.6 1.2 53  TTAG 41.8 41.2 1.3 53  TTAT 40.6 39.8 1.1 51  TTCA 36.0 34.8 1.6 40  TTCC 28.7 28.6 1.3 39  TTCG 33.2 32.5 1.6 62  TTCT 32.4 31.6 1.6 46  TTGA 33.1 32.8 1.2 30  TTGC 27.3 27.6 1.0 45  TTGG 32.1 32.0 1.0 46  TTGG 32.1 32.0 1.0 46  TTGT 28.8 28.8 0.8 55  TTTA 29.2 28.3 1.3 67  TTTC 22.2 21.8 1.1 58  TTGG 25.9 26.2 0.9 63					
TTAC         35.6         35.6         1.2         53           TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63					39
TTAG         41.8         41.2         1.3         53           TTAT         40.6         39.8         1.1         51           TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63					53
TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63					53
TTCA         36.0         34.8         1.6         40           TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63					
TTCC         28.7         28.6         1.3         39           TTCG         33.2         32.5         1.6         62           TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63	TTCA				40
TTCT         32.4         31.6         1.6         46           TTGA         33.1         32.8         1.2         30           TTGC         27.3         27.6         1.0         45           TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63	TTCC				39
TTGA     33.1     32.8     1.2     30       TTGC     27.3     27.6     1.0     45       TTGG     32.1     32.0     1.0     46       TTGT     28.8     28.8     0.8     55       TTTA     29.2     28.3     1.3     67       TTTC     22.2     21.8     1.1     58       TTTG     25.9     26.2     0.9     63	TTCG	33.2	32.5	1.6	62
TTGA     33.1     32.8     1.2     30       TTGC     27.3     27.6     1.0     45       TTGG     32.1     32.0     1.0     46       TTGT     28.8     28.8     0.8     55       TTTA     29.2     28.3     1.3     67       TTTC     22.2     21.8     1.1     58       TTTG     25.9     26.2     0.9     63	TTCT	32.4	31.6	1.6	46
TTGC     27.3     27.6     1.0     45       TTGG     32.1     32.0     1.0     46       TTGT     28.8     28.8     0.8     55       TTTA     29.2     28.3     1.3     67       TTTC     22.2     21.8     1.1     58       TTTG     25.9     26.2     0.9     63		33.1	32.8	1.2	30
TTGG         32.1         32.0         1.0         46           TTGT         28.8         28.8         0.8         55           TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63	TTGC				45
TTTA         29.2         28.3         1.3         67           TTTC         22.2         21.8         1.1         58           TTTG         25.9         26.2         0.9         63		32.1	32.0		46
TTTC 22.2 21.8 1.1 58 TTTG 25.9 26.2 0.9 63	TTGT	28.8	28.8	0.8	55
TTTG 25.9 26.2 0.9 63	TTTA	29.2	28.3	1.3	67
	TTTC	22.2	21.8	1.1	58
TTTT 22.2 22.8 0.9 89	TTTG	25.9	26.2	0.9	63
	TTTT	22.2	22.8	0.9	89

## **References:**

- 20. Needleman SB & Wunsch CD (1970) A General Method Applicable to Search for Similarities in Amino Acid Sequence of 2 Proteins. *Journal of Molecular Biology* 48(3):443-453.
- 21. Durbin R, Eddy S, Krogh A, & Mitchison G (2006) *Biological sequence analysis* pp 92-96.
- Dawson E, *et al.* (2001) A SNP resource for human chromosome 22: extracting dense clusters of SNPs from the genomic sequence. *Genome Res* 11(1):170-178.